

Adapted Convex Optimization Algorithm for Wavelet-Based Dynamic PET Reconstruction

Nelly Pustelnik, Caroline Chaux, Jean-Christophe Pesquet, Florent C. Sureau, Elodie Dusch and Claude Comtat

Abstract—This work deals with Dynamic Positron Emission Tomography (PET) data reconstruction, considering time as an additional variable (space+time). A convex optimization approach closely related to a Bayesian framework is adopted. The objective function to be minimized is expressed in the wavelet-frame domain and is non-necessarily differentiable in order to promote sparsity. We propose an adapted version of Forward-Backward-Douglas-Rachford (FBDR) algorithm to solve the resulting minimization problem. The effectiveness of this approach is shown with simulated dynamic PET data. Comparative results are also provided.

Index Terms—Dynamic PET, wavelet-frame representations, convex optimization, reconstruction.

I. INTRODUCTION

In Positron Emission Tomography (PET), a main challenge consists of finding new reconstruction methods to improve image quality degraded during the acquisition process. Iterative reconstruction methods such as the Expectation Maximization algorithms (which maximize the Poisson log-likelihood associated with PET data) have been proposed to achieve this objective [1], [2]. Accelerated versions of these algorithms have been suggested such as Ordered Subsets EM (OS-EM) [3] and a modified form called RAMLA [4] for which the convergence is established. A drawback of ML-EM approaches is that they converge to noisy images and, in practice, it is thus required to stop iterations before convergence. To overcome this problem, Maximum A Posteriori (MAP) approaches have been developed which take into account a prior controlling noise effects. The major difficulty is to find an appropriate prior. Markov Random Fields (MRF) were first proposed, which are well-adapted to image features. Multiresolution models such as wavelet-based priors were also proved beneficial. A MAP-EM algorithm based on a Gaussian fidelity term and generalized Gaussian distributions to model the wavelets coefficients have been proposed in [5]. The same authors [6] adapted the RAMLA algorithm to solve a MAP estimation problem under the constraint that the regularization term is differentiable (BSREM algorithm [7]).

A new challenge consists of considering dynamic PET data (space+ t) to extract additional physiological parameters. For

dynamic data, the acquisition duration per time-frame is short and thus data noise is more severe. Usually, for dynamic studies, each time-frame is reconstructed independently by using the aforementioned methods. However, in [8], [9], [10] the authors demonstrated the advantage of taking into account time-frame characteristics. These methods can be divided into two classes: on the one hand, methods which directly reconstruct parametric maps [8], [9] and on the other hand, methods reconstructing an image [10], [11], [12]. In parallel, new convex optimization algorithms have been proposed in [13], [14], [15] which are able to solve numerically MAP estimation problems. The main advantages of these algorithms are twofold: *i*) their ability to deal with a wide class of problems and *ii*) guaranty of their convergence. In this paper, we propose to apply one of these appealing methods to space+ t PET reconstruction. Previously, Verhaeghe et al. [10] have suggested to use Forward-Backward-iterations [13] to minimize a criterion with a Gaussian data fidelity term and a wavelet spline regularization. In [11], the authors performed TAC denoising before reconstruction by using the Douglas-Rachford algorithm. In this paper, we propose to develop a nested iterative algorithm (FBDR algorithm) [15] to directly address a Poisson linear degradation model and to perform denoising and reconstruction simultaneously. The convergence proof of this kind of algorithm is given in [15]. This method has another advantage: the possibility to constrain the dynamic range of the image intensity.

This paper is organized as follows. In a first part, we will present the degradation model and the associated objective function to be minimized. The use of multidimensional wavelet representations will be motivated. Then, the FBDR algorithm will be introduced in the case of dynamic reconstruction (space+ t) and the quadratic extension necessary to use this algorithm will be presented next. Finally, we will provide some results for simulated $2D + t$ PET data.

II. PET RECONSTRUCTION MODEL

A. Model

We consider the following degradation model:

$$(\forall t \in \{1, \dots, T\}) \quad z_t = \mathcal{P}(A\bar{y}_t) \quad (1)$$

where $\bar{y}_t = (\bar{y}_{i,t})_{1 \leq i \leq N}$ represents a finite parameterization of the original image and corresponds to the spatial activity distribution for a time-frame t . Here, $z_t = (z_{j,t})_{1 \leq j \leq M}$ is the dynamic PET data corresponding to the number of coincidences for each tube of response for a time-frame t . Finally, A denotes the system linear operator associated with matrix $(A_{j,i})_{1 \leq j \leq M, 1 \leq i \leq N}$ where each element $A_{j,i}$ represents the

N. Pustelnik, C. Chaux and J.-C. Pesquet are with the LIGM, Université Paris-Est – 5, bd Descartes, 77454 Champs sur Marne. Corresponding author: Nelly Pustelnik, E-mail: nelly.pustelnik@univ-paris-est.fr.

F. C. Sureau is with Vrije Universiteit Brussel, Department of Nuclear Medicine, Brussels, Belgium

E. Dusch and C. Comtat are with CEA, DSV, I2BM, SHFJ, Orsay, F-91401, France

This work was supported by the Agence Nationale de la Recherche under grant ANR-05-MMSA-0014-01

probability for a voxel i to be detected in tube of response j . During the acquisition process, data are contaminated by a Poisson noise. The effect of the noise is denoted by \mathcal{P} .

B. Maximum A Posteriori and Convex Optimization

Following a Bayesian approach, the MAP estimate can be expressed as

$$\hat{y}_{\text{MAP}} = \arg \max_y \{ \log p(z|y) + \log p(y) \} \quad (2)$$

where, on the one hand, $p(z|y)$ corresponds to the Poisson likelihood such that $p(z|y) = \prod_{t=1}^T \prod_{j=1}^M p(z_{j,t}|u_{j,t})$ with $u_{j,t} = \sum_{i=1}^N A_{ji} y_{i,t}$,

$$p(z_{j,t}|u_{j,t}) = \frac{(u_{j,t})^{z_{j,t}}}{z_{j,t}!} \exp(-u_{j,t}) \quad (3)$$

and on the other hand, $p(y)$ is an *a priori* distribution on the original image, which will be assumed to correspond to a log-concave function.

In many convex optimization problems, the goal is similarly to minimize a convex objective function by finding

$$\hat{y}_{\text{CO}} = \arg \min_{y \in \mathbb{R}^{N \times T}} \tilde{g}(y) + \tilde{f}(y) \quad (4)$$

where \tilde{f} and \tilde{g} are functions in the class $\Gamma_0(\mathbb{R}^{N \times T})$ of lower semicontinuous convex functions taking their values in $] -\infty, +\infty]$ which are proper (i.e. not identically equal to $+\infty$) and defined on the Hilbert space $\mathbb{R}^{N \times T}$. In the context of inverse problems, \tilde{g} is considered as the data fidelity term and \tilde{f} is a regularizer. Actually, a link between (2) and (4) can be established [16] by identifying \tilde{g} with the Poisson antilog-likelihood term and \tilde{f} with the prior potential function.

Then, the main difficulty lies in the choice of the prior. In the last decade, wavelet representations have demonstrated their efficiency in modelling prior information for images involved in general inverse problems and in PET particularly [6], [10], [5]. In the following, the minimization problem (4) is reformulated in the wavelet transform domain.

C. Wavelets representation and redundant representation

Let $F^* : \mathbb{R}^K \rightarrow \mathbb{R}^{N \times T}$ represent a tight wavelet-frame synthesis operator such that $F^* \circ F = \nu \text{Id}$ with $\nu \in]0, +\infty[$ and let $x \in \mathbb{R}^K$ denote the wavelet coefficients of a field $y \in \mathbb{R}^{N \times T}$ [16]. The field \bar{y} can be expressed as $\bar{y} = F^* x = (F_t^* \bar{x})_{1 \leq t \leq T}$ and Model (1) becomes:

$$z_t = \mathcal{P}(AF_t^* \bar{x}) \quad (5)$$

where \bar{x} represents the unknown vector of wavelet-frame coefficients. In this context, the minimization problem (4) can be reformulated as

$$\min_{x \in \mathbb{R}^K} g(x) + f(x) \quad (6)$$

where the noise related term g is such that $g = \sum_{t=1}^T g_t$ where $g_t = \Psi_t \circ A \circ F_t^*$ and, for every $u_t = (u_{j,t})_{1 \leq j \leq M} \in \mathbb{R}^M$,

$$\Psi_t(u_t) = \sum_{j=1}^M \psi_{j,t}(u_{j,t}) \quad (7)$$

with $\psi_{j,t}(u_{j,t}) = -\log p(z_{j,t}|u_{j,t})$, $j \in \{1, \dots, M\}$. For the regularization term, previous studies have emphasized the advantage of taking the l_0 “norm” as a penalty term so as to promote sparsity but the main difficulty is the non-convexity of this “norm”. The l_1 -norm was then proposed as a penalty term in [13]. Good performance was demonstrated in spite of the possible non-uniqueness of the solution. In [16], authors proposed to take a weighted combination of an l_1 -norm and an l_p -norm with $p > 1$ to regularize the solution, so ensuring the uniqueness of the solution and a more accurate modelling of wavelet-frame coefficients. In this work, we will consider the latter regularizer. Furthermore, a positivity constraint on the image is introduced. More generally, let C be a closed convex in \mathbb{R}^K allowing us to take into account the image range values, then the minimization problem is reformulated as:

$$\min_{x \in \mathbb{R}^K} g(x) + f(x) + \iota_C(x) \quad (8)$$

where ι_C corresponds to the indicator function of the set C , such that $(\forall x \in \mathbb{R}^K)$, $\iota_C(x) = 0$ if $x \in C$ and $\iota_C(x) = +\infty$ otherwise.

III. FBDR ALGORITHM

To solve Problem (8), a nested iterative algorithm named Forward-Backward-Douglas-Rachford (FBDR) was proposed in [15], which consists of a Douglas-Rachford [14] inner loop in each Forward-Backward iteration [17]. To guarantee the convergence to a solution to Problem (8), a necessary condition is the β -Lipschitz differentiability of the fidelity term which is not guaranteed for g as defined in (7). A quadratic extension was proposed in [15] to circumvent this problem.

A. Quadratic extension

The idea behind the quadratic extension approach is to notice that the following inequality is satisfied for the second derivative of function $\psi_{j,t}$ with $t \in \{1, \dots, T\}$ and $j \in \{1, \dots, M\}$:

$$(\forall v \in]0, +\infty[) 0 \leq \psi_{j,t}''(v) \leq \theta \Leftrightarrow v \geq v_{j,t}(\theta) = (z_{j,t}/\theta)^{1/2} \quad (9)$$

when $\theta \in]0, +\infty[$. It is then possible to build a lower approximation of $\psi_{j,t}$ that is θ -Lipschitz differentiable, denoted by $\psi_{\theta,j,t}$. For every value over threshold $v_{j,t}(\theta)$, $\psi_{\theta,j,t}$ is chosen equal to the Poisson antilog-likelihood, whereas for values lower than the threshold, $\psi_{\theta,j,t}$ takes a quadratic form. A lower approximation $g_\theta = \sum_{t=1}^T g_{\theta,t}$ with $g_{\theta,t} = \Psi_{\theta,t} \circ A \circ F_t^*$ is then obtained, which is $(\nu \|A\|^2 \theta)$ -Lipschitz differentiable, where $(\forall u_t = (u_{j,t})_{1 \leq j \leq M} \in \mathbb{R}^M)$, $\Psi_{\theta,t}(u_t) = \sum_{j=1}^M \psi_{\theta,j,t}(u_{j,t})$ and $(\forall v \in \mathbb{R})$, if $z_{j,t} > 0$

$$\psi_{\theta,j,t}(v) = \begin{cases} v - z_{j,t} + z_{j,t} \ln \left(\frac{z_{j,t}}{v} \right) & \text{if } v \geq v_{j,t}(\theta) \\ \frac{\theta}{2} v^2 + \zeta_{j,t,1}(\theta) v + \zeta_{j,t,0}(\theta) & \text{if } 0 \leq v < v_{j,t}(\theta) \\ +\infty & \text{otherwise} \end{cases} \quad (10)$$

and, if $z_{j,t} = 0$, $\psi_{\theta,j,t}(v) = v$, if $v \geq 0$, and $+\infty$, otherwise.

Note that a different polynomial approximation of the objective function was considered in [18].

B. Application of FBDR to space+t PET reconstruction

The proposed reconstruction algorithm possesses some appealing features. At first, the gradient descent is performed in the space domain based on a time-frame by time-frame computation. Secondly, it is grounded on an adapted wavelet decomposition performed along the space/time dimensions taking into account the fewer number of available samples in time.

Algorithm III.1

- ① Choose sequences $(\gamma_n)_{n \in \mathbb{N}}$ and $(\lambda_n)_{n \in \mathbb{N}}$ such that $\gamma_n \in]0, 2/(\nu\|A\|^2\theta)[$ and $\lambda_n \in]0, 1[$. Set $\tau \in]0, 2[$.
- ② Set $n = 0$ and $x^{(0)} \in C$.
- ③ Compute $y^{(n)} = F^*x^{(n)}$.
- ④ For each time-frame $t \in \{1, \dots, T\}$, compute $d_t^{(n)} = A^* \nabla \Psi_{\theta,t}(Ay_t^{(n)})$.
- ⑤ Set $p^{(n)} = x^{(n)} - \gamma_n F d^{(n)}$.
- ⑥ Set $z^{(n,0)} = 2 \text{prox}_{\gamma_n f} p^{(n)} - p^{(n)}$.
- ⑦ For $m = 0, \dots, M_n - 1$
 - a) Compute $z^{(n,m+\frac{1}{2})} = P_C \left(\frac{z^{(n,m)} + p^{(n)}}{2} \right)$.
 - b) Choose $\tau_{n,m} \in [\tau, 2]$.
 - c) Compute $z^{(n,m+1)} = z^{(n,m)} + \tau_{n,m} (\text{prox}_{\gamma_n f}(2z^{(n,m+\frac{1}{2})} - z^{(n,m)}) - z^{(n,m+\frac{1}{2})})$.
 - d) If $z^{(n,m+1)} = z^{(n,m)}$, then goto ⑧.
- ⑧ Set $x^{(n+1)} = x^{(n)} + \lambda_n (z^{(n,m+\frac{1}{2})} - x^{(n)})$.
- ⑨ Increment n and goto ③.

In Step ①, the step-size and relaxation parameters used in the Forward-Backward iterations are chosen. Step ② initializes the wavelet coefficient vector to an element of the convex set C . Step ③ to ⑤ correspond to the gradient descent. The gradient of g_θ is defined as: $(\forall x \in \mathbb{R}^K) \nabla g_\theta(x) = F(A^* \nabla \Psi_{\theta,t}(AF_t^*x))_{1 \leq t \leq T}$ where $(\forall u_t = (u_{j,t})_{1 \leq j \leq M} \in \mathbb{R}^M)$,

$$\nabla \Psi_{\theta,t}(u_t) = (\psi'_{\theta,j,t}(u_{j,t}))_{1 \leq j \leq M}, \quad (11)$$

$$\psi'_{\theta,j,t}(u_{j,t}) = \begin{cases} 1 - \frac{z_{j,t}}{u_{j,t}} & \text{if } z_{j,t} > 0 \text{ and } u_{j,t} \geq v_{j,t}(\theta) \\ \theta u_{j,t} + \zeta_{j,t,1}(\theta) & \text{if } z_{j,t} > 0 \text{ and } 0 \leq u_{j,t} < v_{j,t}(\theta) \\ 1 & \text{if } z_{j,t} = 0 \text{ and } u_{j,t} \geq 0. \end{cases} \quad (12)$$

Step ⑦ allows us to compute the proximity operator of $\iota_C + f$ at point $p^{(n)}$. Details on proximity operators can be found in [17], [14], [16]. The corresponding Douglas-Rachford procedure is mainly decomposed in two operations performed iteratively: on the one hand, a projection P_C onto the convex set C , and on the other hand, the computation of $\text{prox}_{\gamma_n f}$. In practice, C is defined from a convex set $C^* \subset \mathbb{R}^{N \times T}$ allowing us to incorporate constraints on the image range values, the positivity constraint, in particular. The relation between the two convex sets is $C = \{x \in \mathbb{R}^K | F^*x \in C^*\}$.

To compute P_C , we use the following relation [14]:

$$P_C(x) = x + \frac{1}{\nu} F(P_{C^*}(F^*x) - F^*x) \quad (13)$$

For the computation of $\text{prox}_{\gamma_n f}$, explicit forms are given in [16]. Step ⑥ represents the initialization of Douglas-Rachford iterations and allows us to ensure its convergence in one

iteration when the proximity operator of f at $p^{(n)}$ belongs to the convex set C [15]. Finally, Step ⑧ corresponds to the relaxation part of the Forward-Backward algorithm.

IV. MATERIALS AND METHODS

Results are presented on two different slices of the *Zubal* brain phantom including two additional arteries. Each $2D + t$ phantom consists of 256×256 voxels. The generated activity corresponds to a [18F]-FDG exam which was simulated and divided in 16 time-frames with a duration varying between 50 seconds for the first time-frames to 5 minutes for the last ones. 288 (radial) \times 144 (angles) sinograms with a radial sampling of 2.247 mm were simulated by analytically projecting the phantom in the presence of Poisson noise. Attenuation, random and scattered coincidences were not simulated. The number of events in Slice 1 varies from 3 for first time-frame to 647162 for the last time-frame. For Slice 2 the event number varies from 47 to 331348.

The FBDR algorithm is run over 2000 iterations. The chosen separable orthonormal wavelets correspond to Daubechies filters of length 6 on 2 resolution levels for spatial decomposition and 1 level of Daubechies-6 on the interval [19] for temporal decomposition. The latter choice is motivated by the small number of time-frames. The parameter θ is chosen equal to 10^{-4} and the parameters associated with the prior have been determined by a maximum likelihood approach. We compare this algorithm with EM stopped at the iteration that gives the lowest MSE (here 10 iterations for Slice 1 and 2). EM with post-reconstruction smoothing using a Gaussian kernel adapted so as to minimize the MSE was also employed. The EM iteration number in this approach is 120 and 100 for Slice 1 and 2, respectively, and a 4.7 mm \times 4.7 mm full width at half maximum (FWHM) Gaussian filter is used.

V. RESULTS

Figs. 1, 2, 3 and 4 display the different reconstruction results for the 4th and 14th time-frames of each slice. As it can be observed, structures are better recovered using the proposed approach. The advantage of taking into account the temporal aspect can be mainly observed for the first time-frames where the noise level is relatively important (Figs. 1 and 3).

Temporal Activity Curves (TAC) are presented in Figs. 5 and 6 for two neighbouring voxels in the cortex (green area in Fig. 1(a)) and arteries (red area in Fig. 3(a)), respectively. The MSE values presented in Table I show the differences between the reconstructed TAC and the original ones. In each situation, the FBDR approach is the more accurate.

VI. CONCLUSION

We have proposed to employ the FBDR algorithm to restore space+time PET data. This method is very flexible. It allows us to consider non-necessarily differentiable priors, which can be of main interest when using wavelet-frame representations. The obtained results for $2D + t$ PET simulations are very encouraging, and we are now working on $3D + t$ data.

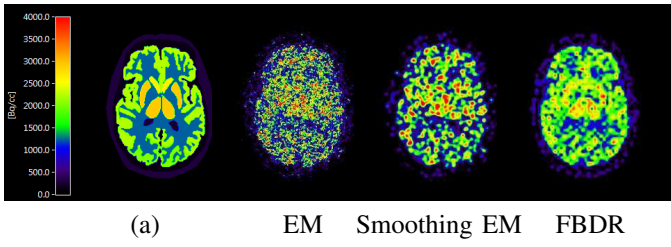


Fig. 1. Slice 1 - time-frame 4.

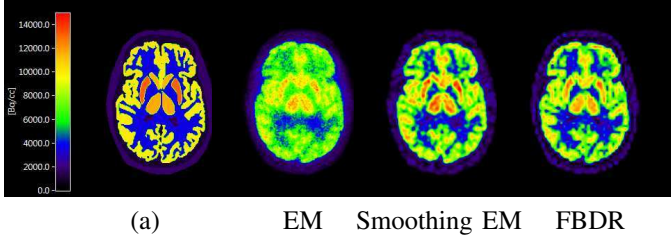


Fig. 2. Slice 1 - time-frame 14.

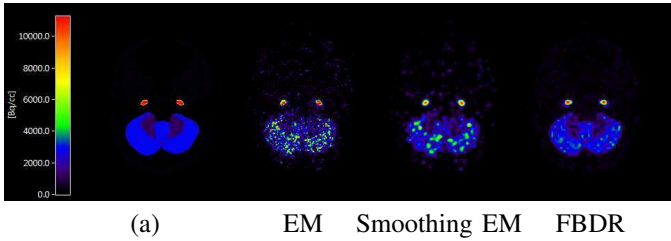


Fig. 3. Slice 2 - time-frame 4.

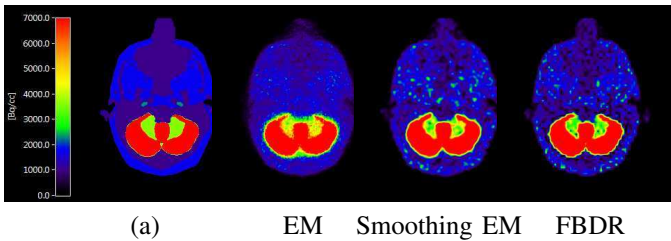


Fig. 4. Slice 2 - time-frame 14.

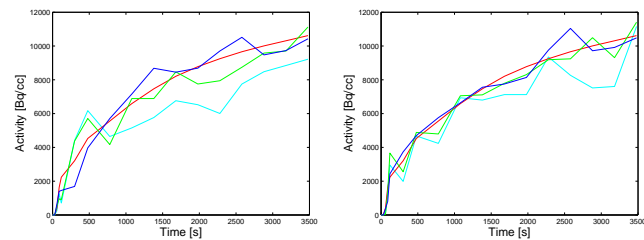


Fig. 5. TAC extracted from Slice 1 for two neighbouring voxels. Original (red), EM (cyan), Smoothed-EM (green) and FBDR (blue).

REFERENCES

- [1] Shepp, L.A., Vardi, Y.: Maximum Likelihood Reconstruction in Positron Emission Tomography. *IEEE Trans. on Med. Imag.* **1**(2) (1982) 113–122
- [2] Lange, K., Carson, R.: EM reconstruction algorithms for emission and transmission tomography. *J. Comput. Assist. Tomogr.* **8**(2) (1984) 306–316

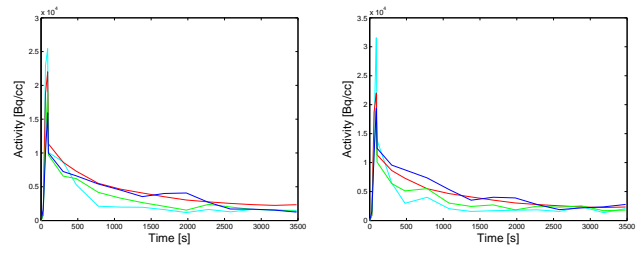


Fig. 6. TAC extracted from Slice 2 for two neighbouring voxels. Original (red), EM (cyan), Smoothed-EM (green) and FBDR (blue).

	Fig. 5 (left)	Fig. 5 (right)	Fig. 6 (left)	Fig. 6 (right)
EM	21.9e+06	16.4e+06	14.9e+06	15.2e+06
Smoothed-EM	4.09e+06	2.45e+06	7.68e+06	5.88e+06
FBDR	2.24e+06	2.20e+06	5.96e+06	3.67e+06

TABLE I
MSE ASSOCIATED TO TAC REPRESENTED IN FIG. 5 AND FIG. 6.

- [3] Hudson, H.M., Larkin, R.S.: Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans. on Med. Imag.* **13**(4) (1994) 601–609
- [4] Browne, J., De Pierro, A.R.: A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography. *IEEE Trans. on Med. Imag.* **15**(5) (1996) 687–699
- [5] Zhou, J., Coatrieux, J.L., Bousse, A., Shu, H.Z., Luo, L.M.: A bayesian MAP-EM algorithm for PET image reconstruction using wavelet transform. *IEEE Trans. on Nuclear Science* **54**(5) (Oct. 2007) 1660–1669
- [6] Zhou, J., Senhadji, L., Coatrieux, J.L., Luo, L.: Iterative PET image reconstruction using translation invariant wavelet transform. *IEEE Trans. on Nuclear Science* **56**(1) (Feb. 2009) 116–128
- [7] De Pierro, A., Yamagishi, M.: Fast EM-like methods for maximum “a posteriori” estimates in emission tomography. *IEEE Trans. on Med. Imag.* **20**(4) (2001) 280–288
- [8] Nichols, T.E., Qi, J., Asma, E., Leahy, R.M.: Spatiotemporal reconstruction of list-mode PET data. *IEEE Trans. on Med. Imag.* **21**(4) (2002)
- [9] Kamasac, M.E., Bouman, C.A., Morris, E.D., Sauer, K.: Direct reconstruction of kinetic parameter images from dynamic pet data. *IEEE Trans. on Med. Imag.* **24** (2005) 636–650
- [10] Verhaeghe, J., Van De Ville, D., Khalidov, I., D’Asseler, Y., Lemahieu, I., Unser, M.: Dynamic PET reconstruction using wavelet regularization with adapted basis functions. *IEEE Trans. on Med. Imag.* **27**(7) (July 2008) 943–959
- [11] Sureau, F.C., Pesquet, J.C., Chaux, C., Pustelnik, N., Reader, A.J., Comtat, C., Trebessen, R.: Temporal wavelet denoising of PET sinograms and images. In: *IEEE Nuclear Science Symp. Conf. Rec.*, Dresden, Germany (19–25 Oct. 2008)
- [12] Reader, A.J., Sureau, F.C., Comtat, C., Trebessen, R., Buvat, I.: Joint estimation of dynamic PET images and temporal basis functions using fully 4D ML-EM. *Phys. Med. Biol.* (2006)
- [13] Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with sparsity constraints. *Comm. Pure Applied Math.* **57** (2004) 1413–1457
- [14] Combettes, P.L., Pesquet, J.C.: A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE Journal of Selected Topics in Signal Processing* **1**(4) (December 2007) 564–574
- [15] Chaux, C., Pesquet, J.C., Pustelnik, N.: Nested iterative algorithms for convex constrained image recovery problems. In: *SIAM Journal on Imaging Sciences*. (2009) to appear.
- [16] Chaux, C., Combettes, P.L., Pesquet, J.C., Wajs, V.R.: A variational formulation for frame-based inverse problems. *Inverse Problems* **23** (June 2007) 1495–1518
- [17] Combettes, P., Wajs, V.R.: Signal recovery by proximal forward-backward splitting. *Multiscale Model. Simul.* **4** (2005) 1168–1200
- [18] Fessler, J.A.: Hybrid poisson/polynomial objective functions for tomographic image reconstruction from transmission scans. *IEEE Trans. on Image Proc.* **4** (1995) 1439–1450
- [19] Cohen, A., Daubechies, I., Vial, P.: Wavelets on interval and fast wavelet transforms. *Appl. and Comp. Harmonic Analysis* (1993) 54–81